R E P O R T R E S U M E S

ED 003 063 24

COMPARISON OF TWO AUTOMATED TEACHING PROCEDURES FOR RETARDED CHILDREN.

BY- HALPASS, LESLIE F. AND OTHERS
UNIVERSITY OF SOUTH FLORIDA, TAMPA

REPORT NUMBER CRP-1267 PUB DATE 63
EDRS PRICE MF-$0.18 HC-$4.24 106P.

DESCRIPTORS- *AUTOINSTRUCTIONAL AIDS, TEACHING MACHINES, *PROGRAM EVALUATION, *PROGRAMED INSTRUCTION, *EDUCABLE MENTALLY HANDICAPPED, *RETARDED CHILDREN, SLOW LEARNERS, MENTAL RETARDATION, LANGUAGE ARTS, WORD RECOGNITION, SPELLING, READING ABILITY, HILLSBOROUGH COUNTY, FLORIDA, PINELLAS COUNTY, TAMPA, GATES WORD RECOGNITION TEST, STANFORD BINET INTELLIGENCE SCALE

A STUDY WAS DESIGNED TO EVALUATE THE USEFULNESS OF AUTOMATED TEACHING PROCEDURES FOR HELPING MENTALLY RETARDED CHILDREN LEARN WORD RECOGNITION, READING, AND SPELLING. THE SUBJECTS FOR THE STUDY (66) WERE DRAWN FROM ESTABLISHED PUBLIC SCHOOL CLASSES FOR THE EDUCABLE MENTALLY RETARDED. SUBJECTS, RANGING IN AGE FROM 8-16 YEARS, WERE MATCHED AND ASSIGNED RANDOMLY TO EITHER AN AUTOMATED TEACHING GROUP OR TO A CONVENTIONAL CLASSROOM GROUP. THREE HYPOTHESES WERE TESTED--(1) THERE WILL BE NO DIFFERENCES BETWEEN AUTOMATED TEACHING AND STANDARD INSTRUCTION FOR TEACHING SELECTED TASKS, (2) EFFECTIVE RETENTION OF SKILLS, TAUGHT THROUGH AUTOMATED MEANS, WILL BE DEMONSTRATED BY POST-LEARNING TASKS, AND (3) NO SIGNIFICANT DIFFERENCES WILL BE FOUND BETWEEN AN AUTOMATED PROCEDURE UTILIZING A MULTIPLE-CHOICE METHOD AND ONE UTILIZING A TYPEWRITER KEYBOARD METHOD. HYPOTHESIS ONE WAS REJECTED IN LIGHT OF SIGNIFICANT SKILL IMPROVEMENT BY SUBJECTS USING BOTH AUTOMATED TEACHING PROCEDURES OVER CONVENTIONAL METHODS. HYPOTHESIS TWO WAS PARTIALLY SUPPORTED. SPELLING IMPROVEMENT WAS RETAINED OVER A RELATIVELY SHORT PERIOD. HYPOTHESIS THREE WAS PARTIALLY SUPPORTED. IN VIEW OF THE FINDINGS, FURTHER RESEARCH CONCERNING AUTOMATED TEACHING WITH THE RETARDED WAS RECOMMENDED. (JC)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

COMPARISON OF TWO AUTOMATED TEACHING PROCEDURES FOR RETARDED CHILDREN

Cooperative Research Project No. 1267

Leslie F. Malpass - Project Director

Alden S. Gilmore - Principal Investigator
Miles W. Hardy - Principal Investigator
Charles F. Williams - Principal Investigator

University of South Florida

Tampa, Florida

1963

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education, and Welfare.

ED003063

ACKNOWLEDGEMENTS

In a study as extensive as this one, the support and co-operation of many agencies and individuals is required. The investigators wish to acknowledge, with sincere appreciation, the assistance given by the following:

- The Cooperative Research Program of the U. S. Office of Education, for underwriting the major financial obligations of the study and for providing valuable guidance.

- The University of South Florida (President John S. Allen) and the MacDonald Training Center (Fischer S. Black, President of MacDonald Training Center Foundation, and Calvin M. Pinkard, Director of Research) for generous allotments of released time, use of facilities, and supplementary funds for equipment and operating expenses necessary for the project.

- The Hillsborough County, Pinellas County public school systems (Crockett Farnell and Floyd Christian, superintendents of schools, respectively). The directors of special education (Minnie Lee Powell and Paul Stevens, respectively) were especially helpful in obtaining the cooperation of the several school principals and thirty-eight teachers from whose classes the public school subjects were selected. The principals were extremely kind in providing adequate space and work schedules for the project.

- The Sunland Training School, Ft. Myers, Florida (J. M. Pressley, superintendent, and Maurice Dayan, supervising

psychologist) graciously permitted use of their facilities and encouraged the children selected for the study to participate in the study. Joel McLean, psychologist, was especially helpful in supervising the sample selection and data collection at the Sunland Training School.

- Barbara Sparkman, Ann Gilmore, Andrée Holland, and Dorothy Ann Kerber who served as research assistants for the study. By carefully administering the automated instructional programs, by assisting with other data collection, and by helping with the tabulation and running of the data for the final report they added significantly to the study.

- Inez Black, Fran Hill, and Marie Clark, who served as part-time secretaries to the project. They assisted with program preparation, maintained necessary records, and performed a myriad of clerical functions in their contributions to the investigation.

- James Holland (Harvard), Lawrence Stolurow (University of Illinois), Richard Anderson (University of Florida), and Clarence Williams (University of Rochester) served as professional consultants. C. J. Donnelly (Teaching Materials Corporation), D. E. Cornell (Teaching Machines, Inc.), and Charles Lockyer (Publishers Company, Inc.) served as technical consultants about teaching machines used in the study.

To all of these go the hearty thanks of the four principal investigators.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		i
LIST OF TABLES		iv
LIST OF ILLUSTRATIONS		v
Chapter		
I	STATEMENT OF THE PROBLEM	1
	Purposes of the Study	6
II	PROCEDURE	9
	Sample	9
	Evaluation Measures	15
	Other Measures	18
	Method	19
	Development of Teaching Machine Programs	23
III	RESULTS	35
	Analysis of Retention	45
	Post-Instruction Retention	51
	Relationships Between Instructional Condition and Other Variables	52
	Reading in Context	57
	Behavioral Ratings	58
IV	CONCLUSIONS AND COMMENTS	61
	Variables Relevant to Automated Instruction	66
APPENDICES		73
REFERENCES		93

LIST OF TABLES

Table		Page
1	Matching Variables - Public Schools Sample	11
2	Matching Variables - Sunland Training Center	14
3	Word Gains Following Instruction - Public Schools Sample	37
4	Word Gains Following Instruction - Sunland Training Center	40
5	Comparison of Matched Group Performance - Mean Word Gains	43
6	Retention Ratios Between Pre-Test to Post-Test Gains and Pre-Test to 30 Day Post-Test Gains for Special Matched Groups	45
7	Relationships Between Programed Word Gain and Related Variables A. Public School Groups	53
8	Relationships Between Programed Word Gain and Related Variables B. Institutional Group	54
9	Relationships Between Programed Spelling Word Gain and Related Variables A. Public Schools Group	56
10	Relationships Between Programed Spelling Word Gains and Related Variables B. Institutional Group	57
11	Mean Behavioral Rating for Subjects With Highest and Lowest Programed Word Gain	59

LIST OF ILLUSTRATIONS

Figure		Page
1	The multiple choice method - The Teachall	21
2	The typewriter keyboard method - The Wyckoff Film Tutor	21
3	Multiple choice program - Sample Frames	27
4	Keyboard method - Sample Frames	29

CHAPTER I

STATEMENT OF THE PROBLEM

The problem of how to teach reading and spelling to retarded children has been of concern to educators for many years. Conventional methods of teaching reading have been explored, and there is some disagreement about the most effective means of presenting materials to be learned. Some teachers use the alphabet approach; others begin with phonics. Both of these methods are primarily oral. They tend to promote word analysis and word recognition skills. They also tend to present difficulties, especially for retarded children, because of the vagaries of the English language itself (Anderson & Dearborn, 1952).

The word method has been cited as the quickest way to teach basic sight vocabulary (op. cit., p. 256). This method presents words as units, and the child learns to recognize combinations of letters as representing objects and events. Many authorities suggest that word recognition is the most efficient way of promoting the vital function of retention (Monroe, 1932; Thorndike, 1934; Cruickshank & Johnson, 1958). Certainly for retarded children, the word method is more suitable than the phrase sentence methods, the strictly phonetic, or the alphabet approaches.

Retention constitutes one of the most trying problems for teachers of retarded children. Many educable retarded children[1]

---

[1]As used in this report, educable mentally retarded (EMR) refers to a person in the IQ range 50 - 80, who is at least two years retarded in school achievement and who has not demonstrated sufficient ability to profit from ordinary classroom instruction.

can learn to recognize word forms au momente. It is characteristic of them, however, not to be able to demonstrate even simple word recognition over a period of time (Kirk, 1940). Conventional procedures, therefore, do not seem to have been particularly effective in teaching retarded children how to read. A survey of 128 educable mentally retarded subjects, 15 and 16 years old who had been enrolled in public school for at least seven years and in special classes for at least the three years immediately prior to the study, indicated that 74 of them read below the third grade level and 97 below the fourth grade level. Only one 16 year old approached the seventh grade level (Powell, 1961). This suggests that many educable retarded children profit minimally from conventional classroom procedures. If some means could be found to enable these children to learn to read, the profit to society would be enormous. This challenge provided the basis for the present study.

Independent observations by the four principal investigators of this study had suggested that the processes for teaching retarded children how to read could be facilitated if the conditions for learning were more appropriate than those ordinarily supplied in the classroom. All of the investigators have worked with retarded children in individual and group situations and have acted as consultants to special education teachers and supervisors in public schools. Prior research by one of them (Malpass, with Carrier & Orton, 1960) had demonstrated that retarded children can learn to manipulate mechanical learning devices as competently as normal children, although their rate of learning was not as high. Another study by Gilmore & Williams (with Pinkard & Ricker, 1963) has shown that ability to read is a factor in the prediction of employability of the retarded.

By 1960, several reports had been published which suggested that automated instructional devices were efficient and economical in

helping normal children at almost every age level to learn various academic skills. Skinner and Porter at Harvard University were experimenting with multiple choice teaching machines; Omar K. Moore at Yale had developed a typewriter keyboard machine that helped two and three year old normal children to learn to read and spell; and Lawrence Stolurow at the University of Illinois was beginning to program arithmetic for retarded children. Teaching machines and automated instruction seemed to offer the possibility of helping retarded children to learn to read and spell, but there was no research available about the utility of the methodology with this group.

Although automated teaching devices have been available for use in American education since 1915 (Ordahl & Ordahl, 1915), it was not until 1927 that Pressey developed machines which presented information, maintained a record of performance, and permitted feedback of response to the learner (Pressey, 1926, 1927). One of Pressey's devices consisted of a punchboard which revealed a color coded spot when the correct answer was chosen. The other was a multiple key apparatus that permitted the learner to progress with programed material only when he gave the correct answer.

Pressey's ideas did not generate much support from educators at the time. It was not until after the publication of B. F. Skinner's article in Science (1958) that psychologists and educators became interested in the potentials of automated instruction. Many people now recognize Skinner's article as a landmark in the history of education. The historical development of this method has been clearly presented elsewhere (e.g., Galanter, 1959; Green, 1962; Lumsdaine & Glaser, 1960; and Deterline, 1962); consequently, it is not necessary to discuss Skinner's contributions here. Suffice to say that concepts of stimulus discrimination and reinforcement provide the primary basis for the experimental analysis of teaching reading and spelling to retarded children that underlie this study.

By 1960, there was sufficient interest in teaching machines and programed instruction to warrant the formation of over thirty companies to produce machines. At least one journal had appeared (The Automated Teaching Bulletin, published by the Rheem Califone Company), along with many articles and two major books. Private foundations and the federal government were beginning to provide support for research activities. Programs had been written for all levels of schools and college students, and for industrial and military personnel (Galanter, 1959). Subjects which had been programed included English, psychology, foreign languages, statistics, physics, electronics, and several elementary level subjects. Although advanced elementary reading and arithmetic programs were being developed, none was available for teaching word recognition, vocabulary, reading, and spelling to retarded children.

Variations occur in the description of what constitutes a teaching machine, but most authorities agree with Stolurow (1960) that there are certain basic functions which any teaching machine should serve for special education.

> "(1) A teaching machine presents information either as statements and then questions, or as paired sets of abstract symbols to be associated and accompanying incomplete associations for testing purposes. This is the presentation function.
>
> (2) A teaching machine requires a response that is overt and measurable; however, the form of the response and the measure taken of it may vary from machine to machine depending upon the educational and/or training purpose involved. This is the response function.
>
> (3) A teaching machine compares, or permits the learner to compare, the response with a predetermined correct response (criterion). This is the comparator function.

(4) A teaching machine provides knowledge of results and has the potential for doing this after every response if it is not, in fact, used that way. This is the feedback function.

(5) A teaching machine programs information (e.g., prose statements, questions, and/or abstract associable symbols or signs) according to a set of rules that are capable of being made explicit if only at the level of being 'point-atable.' This is the programing function.

(6) A teaching machine allows for variability in the information presentation rate. This is the pacing function.

(7) A teaching machine combines learning and its measurement into a single set of coordinated operations. This is the collator function.

(8) A teaching machine automatically rejects the correctly-responded-to items so that the learner sees a smaller number of items from trial to trial. This is the selective function." (pp. 431-32).

Skinner-type multiple choice machines were available from a number of commercial sources. The one selected incorporated most of the desirable data expected of automated teaching devices and supplied both auditory and visual feedback to the learner (Skinner, 1958; Holland, 1960). A typewriter keyboard machine, called the Film Tutor, adapted from Moore's model by Wyckoff (1960) was also available.

In essence then, multiple choice and modified completion-type procedures held specific promise for teaching word recognition, reading, and spelling skills to retarded children. Because machines were available to present these types of procedures, it was decided to compare their relative effectiveness for these purposes.

The feasibility of the study was enhanced by the interest of the staff at the MacDonald Training Center, located in Tampa, Florida.

The Center was established to develop methods for helping retarded individuals gain practical skills. The staff works closely with special educational personnel from the Hillsborough County and Pinellas County public school systems. Personnel from both school systems subsequently offered to cooperate in the research project. Classes for retarded children from these two school systems became sources for the public school subjects used in the study. The Sunland Training Center at Ft. Myers, Florida, also cooperated by supplying groups of institutionalized retarded children as subjects for the automated teaching procedures.

Purposes of the Study

It is the major purpose of this study to evaluate the usefulness of two automated teaching procedures for helping mentally retarded children to learn word recognition, reading, and spelling. One procedure utilizes a multiple choice method adapted from linear programing methods and multiple choice machines devised by Skinner and his colleagues. The other procedure utilizes an automated typewriter keyboard apparatus designed by Wyckoff (1960) based on principles reported by Omar K. Moore (1960).

One major aspect of the study is concerned with determining whether these procedures are effective for teaching word recognition, reading, and spelling to retarded children enrolled in public school special education classes and to institutionalized retardates comparable in intellectual status. Effectiveness in teaching these academic skills is measured principally in terms of acquisition and retention of 72 basic words, used singly and in grammatical context. A comparison of the two automated teaching procedures constitutes the other major aspect of the study.

The general hypotheses of the study are:

1. No significant differences in efficiency will be found between automated teaching procedures and standard classroom instruction procedures for teaching word recognition, reading and spelling to mental retardates.

2. Effective retention of word recognition, reading, and spelling skills obtained in automated teaching procedures will be demonstrated by post-learning tests.

3. No significant differences in learning and retention will be found between an automated teaching procedure utilizing a multiple choice method and one utilizing a typewriter keyboard (modified completion) method.

In brief, the study attempts to evaluate the usefulness of automated teaching procedures in contrast to ordinary teaching methods, and to compare the usefulness of differing automated procedures.

PRECEDING PAGE BLANK-NOT FILMED

CHAPTER II

PROCEDURE

Sample

The subjects for this study were drawn from public schools in Hillsborough County and Pinellas County, Florida. The public school subjects (N=66) constitute the major sample of this study and will be referred to as the "public school group."

The Florida school system, like many others in the country, has established special classes for the educable mentally retarded. The Hillsborough County special education classes include educable youngsters ranging in IQ from 50 to 80. Their ages range from about 8 years to 12 years in the elementary school classes, and from 13 to 16 years in the junior high school classes. "Junior high school classes" mean that some educable classes are located in junior high schools. These classes do not, of course, present junior high school level curriculum materials. In addition to intelligence and age, academic retardation and poor adjustment in classes provide admission criteria for special education classes. There are 26 classes located in 19 schools in the county. The elementary and intermediate classes are limited to an enrollment of 15, and the junior high classes to an enrollment of 18. In Hillsborough County subjects were drawn from 22 of the 26 special education classes.

In Pinellas County the majority of the educable mentally retarded classes are located in one school. Subjects for this study were drawn from eight different classes at the Nina Hartis School in St. Petersburg, Florida.

Early in the study an opportunity was presented to gather data from a state institution for the retarded. With the cooperation of the psychologists and staff at Sunland Training Center, Ft. Myers, Florida, it was possible to establish automated teaching conditions similar to those for the public school group. Although the original proposal did not call for collecting data from a sample of institutionalized retarded children, this provided an excellent opportunity to extend the study; and so data were collected from a group of institutionalized subjects (N=33). These subjects will be referred to as the "institutional group."

Subjects were matched as closely as possible in groups of three on the variables described below. Members of each triad were randomly assigned to one of the following groups:

Experimental Group 1 - Multiple Choice Method
Experimental Group 2 - Keyboard Method
Classroom Group - Conventional special education classroom instruction

Subjects were matched on sex, chronological age, mental age, number of programed words recognized, and on a standardized reading achievement test (Gates, 1958). In addition, an attempt was made to equate socio-economic level and visual and orthopedic handicaps.

Characteristics of the Public School Group

Variables upon which subjects in the public schools sample were matched are summarized in Table 1.

TABLE 1

MATCHING VARIABLES

Public Schools Sample

	Multiple Choice	Keyboard	Classroom
Sex Distribution			
No. Males	16	16	16
No. Females	6	6	6
Chronological Age			
Mean	11.8	12.5	12.0
Range	10.0-15.9	10.0-15.7	10.0-15.6
Mental Age[1]			
Mean	8.2	8.2	8.3
Range	5.7-11.4	6.0-11.0	6.4-10.9
Programed Words Known			
Mean	10.2	14.3	13.4
Range	0-26	0-34	0-38
Gates Word Recognition Test-Grade Placement			
Mean	1.79	1.94	1.74
Range	1.3-2.35	1.3-2.5	1.3-2.8

[1] Mental ages were calculated from scores on the Stanford-Binet Intelligence Scale (1937 revision)

The public schools group was comprised of 48 boys and 18 girls, reflecting the sex distribution of the educable mentally retarded classes from which the sample was drawn. An identical ratio of boys and girls was maintained in each of the three instructional groups.

The chronological age of the public schools group ranged from 10 years to 15.9 years with a mean of 12.1 years. The mean CA for the multiple choice group was 11.8 years, keyboard group 12.5 years, and for the classroom group 12.0 years.

The mental age of the public schools group ranged from 5.7 years to 11.4 years with a mean of 8.2 years. The mean MA for the multiple choice group was 8.2 years, keyboard group 8.2 years, and for the classroom group 8.3 years.

The number of programed words recognized prior to teaching machine instruction ranged from 0 to 34 words for the two experimental groups with a mean of 10.2 for the multiple choice method, and a mean of 14.3 words for the keyboard method. Word recognition by the classroom group ranged from 0 to 38 with a mean of 13.4.

Reading grade placement, measured by the Gates Primary Reading Test - Word Recognition, ranged from the 1.3 year level to the 2.8 year level with a mean reading grade level of 1.8 years. The three subgroups showed mean grade placement scores as follows: multiple choice group 1.79 years, keyboard group 1.94 years, and classroom group 1.72 years.

An attempt was made to equate socio-economic factors in the assignment of subjects in the three subgroups by obtaining teachers' ratings of socio-economic level. Teachers were asked to designate the socio-economic level of the subjects on a three point scale of high, medium, and low socio-economic level for matching purposes.

Health examinations administered by the public school system provided information on visual, orthopedic, and other handicaps. Children with severe handicaps were excluded from the study. In cases where moderate physical handicapping conditions did exist, an attempt was made to control for this factor by equated assignments to the three subgroups.

Because of the care in matching on other criteria and because of the difficulty in assessment of neurological damage, there was no attempt to classify the subjects according to neurological categories. However, where signs of brain damage were evidenced through pre-test behavior and teachers' reports, they were included as matching variables.

Characteristics of the Institutional Group

Variables used to match the institutional group are summarized in Table 2.

Subjects in the institutional group ranged in age from 11 years to 20.6 years, with a mean chronological age of 15.5 years. The mean CA for the multiple choice group was 15.1 years, for the keyboard group 16.0 years, and for the classroom group 15.4 years. The sex distribution was similar to the public schools group.

The mental ages of the institutional group ranged from 6.9 years to 14.5 years, with a mean MA of 9.9 years. The mean MA for the multiple choice group was 10.2 years, for the keyboard group 10.3 years, and for the classroom group 9.1 years.

Reading grade placement, measured by the Gates Primary Reading Test-Word Recognition, ranged from 1.3 years to 2.3 years with a

TABLE 2
MATCHING VARIABLES
Sunland Training Center

	Multiple Choice	Keyboard	Classroom
Sex Distribution			
No. Males	8	8	8
No. Females	3	3	3
Chronological Age			
Mean	15.1	16.0	15.4
Range	11.1-20.6	11.4-19.1	11.0-17.4
Mental Age			
Mean	10.2	10.3	9.1
Range	6.9-14.5	8.0-12.8	7.1-11.4
Programed Words Known			
Mean	12.0	14.2	10.0
Range	2 - 24	7 - 23	1 - 25
Gates Word Recognition Test-Grade Placement			
Mean	1.8	1.8	1.6
Range	1.3-2.3	1.4-2.1	1.3-2.1

mean reading grade level of 1.7 years. The mean reading grade levels of the three subgroups were: multiple choice group 1.8 years, keyboard group 1.8 years, and classroom group 1.6 years.

Pre-test recognition of programed words ranged from 1 to 25 words with a mean score of 12.1 words recognized. The mean word recognition for the three subgroups was as follows: multiple choice group 12.0 words, keyboard group 14.2 words, and classroom group 10.0 words.

Evaluation Measures

The following measures were the dependent variables for the pre-test group. Each measurement was made prior to instruction and immediately upon completion of instruction. Retests were obtained at 30 and 60 days following completion of instruction.

Pre and post-test data were gathered on the following:

(1) Word Recognition - Programed Words
(2) Word Recognition - Non-Programed Words
(3) Spelling - Programed Words
(4) Spelling - Non-Programed Words
(5) Gates Primary Reading Test - Word Recognition

A paragraph reading test and a rating scale were administered upon completion of instruction.

Word Recognition - Programed Words

For testing purposes the 72 programed words were typed on individual 3" x 5" cards. (Criteria for selection of words are

described in the section, "Development of Teaching Machine Programs," also listed in Appendix A.) These 72 word cards were intermixed with the 28 non-programed word cards, and the cards were shuffled in random order for presentation to the subject. Testing for word recognition was administered individually to each subject by a research assistant. Research assistants would call one child at a time out of the classroom into a separate testing room. In all cases the assistant had prior contacts with the subject before the individual testing situation occurred. The research assistant informed the child that the test was not a school test and asked him to look for "hard" and "easy" reading words. Subjects were told they were not expected to know all the words which would be shown them on the cards.

The assistant sat beside the subject at the desk or table and stacked the 100 cards in front of the subject. The cards were turned over in order, and the subject was given three seconds to respond to each word. The scores for right and wrong responses were recorded by the research assistant in such a way that the subjects indicated they were not aware of how their responses were being scored. As subjects responded, the research assistant would stack the cards in overlapping piles with the correct responses on the right and the incorrect responses on the left. After the student left the room, the student's performance would be recorded.

Although some subjects showed evidence of anxiety during the testing situation, it did not appear that this blocked the recognition of highly familiar words. Where doubt existed about the validity, results were checked with the classroom teacher's evaluation of the student's classroom performance. If it were presumed that better performance should be expected, the test would be readministered.

Word Recognition - Non-Programed Words

Twenty-eight words of equivalent difficulty to the programed words were typed on individual 3" x 5" cards and intermixed with the 72 programed words for presentation purposes. (See Appendix A)

Spelling Test - Programed Words

Every fourth word of the 72 programed words made up the programed spelling test. (See Appendix A) Upon completion of the word recognition test, subjects were asked to spell the 18 programed and 7 non-programed spelling words. A research assistant gave the following instructions: "I want to see how many of the following words you are able to spell. I will call them out and you try to spell them as well as you can." Subjects were not told of their scores on the spelling test.

Spelling Test - Non-Programed Words

Every fourth word of the 28 non-programed word list was selected for this test. Eighteen programed words and seven non-programed words made up the spelling test. (See Appendix A). These words were given to subjects for spelling mixed randomly among the 18 programed spelling words.

Gates Primary Reading Test - Word Recognition

The Gates Primary Reading Test - Word Recognition section was used as an estimate of reading grade level. This test is designed to sample ability to read words representative of the primary vocabulary. It consists of 48 frames, each of which contains

four printed words and a picture which illustrates the meaning of one of them. The test was administered according to the standard instructions.

Other Measures

In addition to programed and non-programed word recognition, programed and non-programed spelling, and reading grade level measurements, the following tests were constructed and administered to all public schools subjects.

Paragraph Reading Test

A paragraph reading test, which included 42 of the programed words used in a story context, was designed by the experimenters. (See Appendix B) This test was used as a part of the followup procedure. Its purpose was to measure the subject's ability to recognize the programed words among other words in a sequence presentation offering cues found in normal reading situations.

Behavioral Rating Scale

A behavioral rating scale was constructed by the experimenters to assess behavioral characteristics associated with word recognition ability. A copy is reproduced in Appendix C. Classroom teachers and research assistants filled out a scale on each subject upon completion of teaching machine instruction.

Four descriptive rating steps were prepared under the headings: Personal Adjustment; Work Habits; Motivation; Direction

Following Ability; Aspirations; Attention and Concentration; and Maturity. While reliability was not established independently for the scale, many of the scale steps were taken from instruments previously found reliable in earlier experimentation with retardates at MacDonald Training Center, Tampa, Florida (1959). Instructions requested ratings in the framework of the total EMR population without regard to expectations for normal students.

For the institutional group data were gathered on only word recognition and reading grade level since it was not possible to administer the other measures.

Method

The major purpose of this study was to evaluate the effectiveness of two automated teaching procedures compared with classroom instruction for mentally retarded children. Exploratory investigation led to the conclusion that programed instruction utilizing a mechanical device with emphasis on self-pacing would be the best method of capturing the attention of retarded subjects and reducing some of the problems of usual instructional methods with this population. As indicated in a previous section, a survey of teaching machine technology had indicated that two appropriate methods of program presentation for use with retarded subjects were the multiple choice technique patterned after the work of Skinner and Holland (1960) and a typewriter keyboard presentation based upon the techniques developed by O. K. Moore (1960) as modified by Benjamin Wyckoff (1960).

Multiple Choice Method

The multiple choice teaching machine initially selected for use was a commercially available electronic apparatus which

utilized 8 mm film strips. This was a fully automated desk-size machine in which programed frames were projected upon a ground glass screen. Response selection was made by pushing small, clear, plastic tabs which overlaid the response choices (i.e., a child would point directly at the response item). A correct response automatically advanced the film to the next frame. Although this apparatus seemed to be well suited for our sample population, considerable technical difficulties were encountered in the operation of the equipment; and it became necessary to abandon plans for using this machine in the project. Instead a manually operated multiple choice machine was substituted for the electronic apparatus originally selected for use. (See Figure 1)

A "Teachall" teaching machine was selected as the multiple choice type of equipment to be used in this study. The programed material for this machine is printed on heavy cardboard cards which are stacked in one side of the machine. A slide lever at the top of the machine mechanically moves the cards from the storage compartment to an open display panel situated directly above five mechanical keys. Since the programing used in this project required a selection of only three responses, two of the keys were colored a different color to more clearly distinguish the response keys to be used. The bottom of the heavy cardboard cards are notched so that contact can be made when the correct response is selected by depressing one of the three response keys. A correct choice is signaled by a small green light on the right hand side of the machine and a buzzer tone. An incorrect response produces no signal.

Keyboard Method

The typewriter keyboard method, or keyboard method as it will be referred to hereafter, is Benjamin Wyckoff's adaptation of the

Fig. 1. The multiple choice method - The Teachall

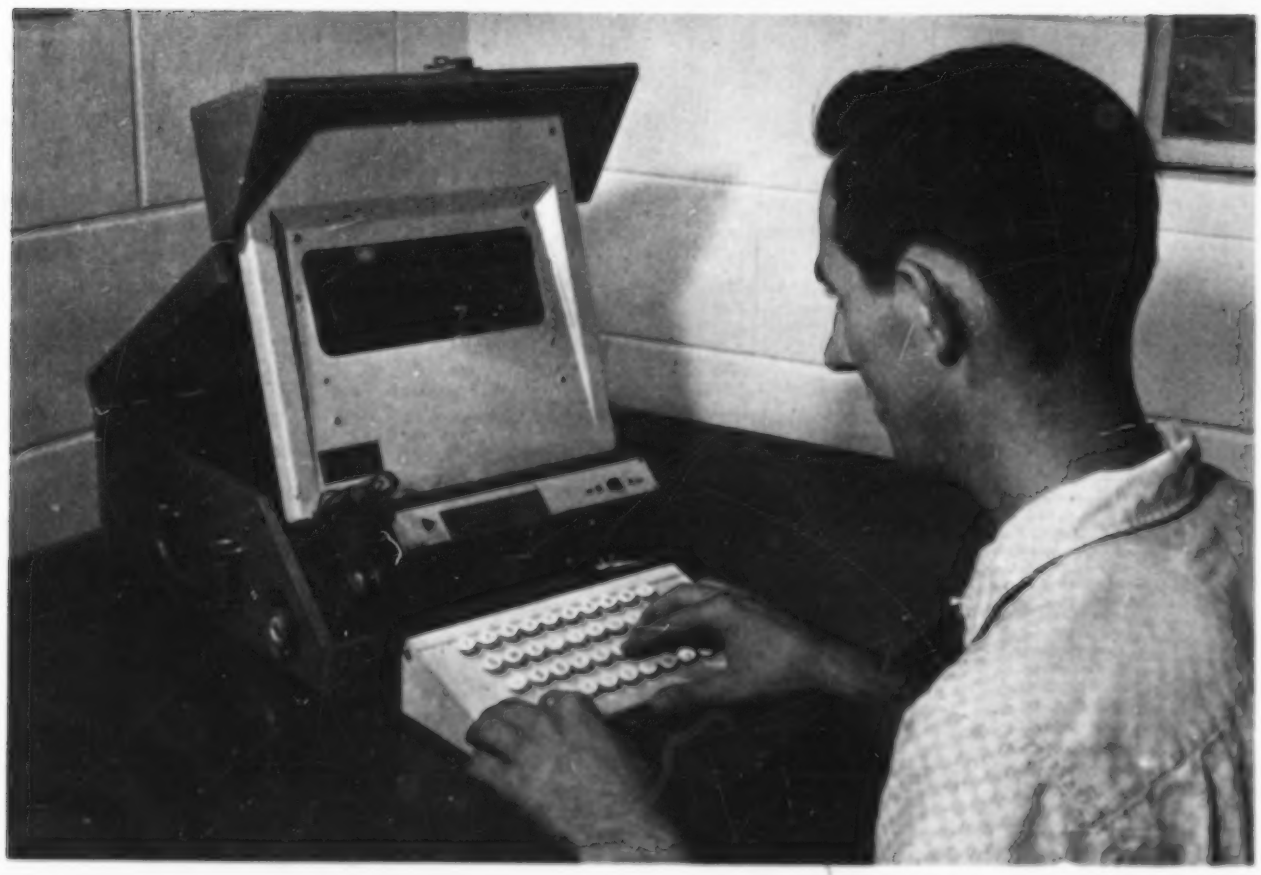

Fig. 2. The typewriter keyboard method - The Wyckoff Film Tutor

Omar Khayyam Moore's original teaching machine research procedure (Moore, 1960). The device, shown in Figure 2, is an electronically actuated projector combined with a typewriter-like keyboard. The machine is loaded with a 35 mm film strip prepared with four frames occupying a normal 35 mm frame space. Each frame, in addition to presenting information and soliciting a response, supplies a coded correct answer to the computer-like circuit which permits the typewriter keyboard to act selectively.

Each frame calls for a specific response in the form of a letter of the alphabet or a number. Only when the subject presses the correct letter or number on the keyboard does the machine respond by advancing to the subsequent frame. At each advance the new frame shows the response (letter or number selected) as though it had been typed by depressing the correct key on a normal typewriter. Discrimination among letters is forced in order for the subject to secure correct responses. Whole word discrimination also is required on some responses.

Classroom Instruction

The two automated teaching methods were compared with conventional classroom instruction. The "classroom group" was composed of students enrolled in 10 special education classes for the educable mentally retarded in Hillsborough County and Pinellas County public schools systems. Teachers were given a list of the 72 programed words presented by the automated techniques and were informed of the nature of the study. They were encouraged to present these words during their regular classroom periods over the two month period covered by the study. Since these words are common to all primary grade level reading material, this presentation did not significantly alter the normal classroom routine. Although no

control was exercised over the teacher's behavior, all of them reported that they had spent at least some time each day presenting the words.

It should be noted that all of the subjects in this study had been exposed to a minimum four years of classroom instruction during which time the subjects had been repeatedly exposed to most of the words included in this study. Very few, if any, of these words could be considered as "new" words to the subjects, even though many of the subjects could not recognize or spell them.

Development of Teaching Machine Programs

At the time of this study no teaching machine program for teaching reading at the primary grade level was available. It was therefore necessary to develop original programs for the two experimental conditions.

Seventy-two words were selected from a basic word list of 365 most used words which appeared in standard elementary reading material (McEathron, 1956). All of the words appear on the Buckingham-Dolch Combined Word List (1936). The 72 words ranged from first through third grade level words in terms of usage frequency. The primary criteria for choosing words were functional applicability and frequency of appearance in standard reading textbooks at the primary grade level. A consideration in the selection of words was their feasibility for pictorial presentation. The illustrative material was field tested for clarity of meaning with retarded subjects not included in this study. In this manner ambiguous illustrations were eliminated in the final program. Programed words included 46 nouns, 10 adjectives, 9 verbs, 4 conjunctions, 2 prepositions, and 1 article. The programs constitute distributed practice in discrimi-

nation and sight recognition and make no use of phonic or oral aids. The 72 programed words and 28 non-programed words used in this study are presented in Appendix A.

Standard programing techniques such as cueing, vanishing, chaining, and fading, as well as direct practice in discrimination and recognition, were used. Nouns and verbs were paired with pictures and generalized with differing pictures. Articles and conjunctions were presented by frequent appropriate insertion without any attempt at explanation.

Early development of the programs was largely by trial and error. Each of the four investigators independently constructed a series of programed frames which were tested on small groups of retarded subjects who were not included in the final population sample. A careful record of error rate was maintained during these pre-study trials, and the programing underwent a series of revisions until an error rate of less than 5 per cent was obtained. In the early steps of validating the two programs frames were made up by hand separately on 4" x 6" cardboard sheets. A research assistant sat beside the subject, turning the cards for him and providing verbal feedback by answering "right" or "wrong" to each response. If the subject did not choose the correct response on the first trial, he was allowed alternative choices until a correct selection was made.

Experience in the validation process indicated that the majority of subjects grew restless if kept at the task for over twenty minutes. It was also found that most subjects could complete 200 frames well within the twenty minutes. Consequently, a daily exposure rate of between 150 and 200 frames per session was anticipated.

Both programs utilized a high degree of repetition. After a new word was introduced into the program, it would be presented

at least thirty times within five consecutive instructional sessions; thereafter it would be reinforced intermittently at a decreasing frequency throughout the remainder of the program. Early validation procedures also showed the most effective rate of word introduction to be two newwords per instructional session if a fixed error rate were to be maintained.

Writing and validation of the two programs developed for this study required several thousand hours of effort. After the programing format was developed, considerable time was required by the technical problems involved with producing the films and cards used for presentation of the final programs to the subjects.

Multiple Choice Program

The final version of the multiple choice program consisted of 3,000 cards, divided into 40 lesson units. Each lesson unit consisted of approximately 75 cards. An average of two new words were introduced in each lesson unit. In the multiple choice program it was possible to expose each unit twice on successive days, thereby allowing duplicate exposure of each frame yielding a total of 6,000 exposures. Pre-test experimentation indicated this degree of repetition did not create resistance among the subjects. Since a large amount of art work was involved in the production of the frames, the repetition factor yielded a considerable savings in costs.

The following overlapping presentation sequence was used:

Session 1 - Lesson Units 1 and 2
Session 2 - Lesson Units 2 and 3
Session 3 - Lesson Units 3 and 4
Session 4 - Lesson Units 4 and 5
Etc.

A representative sample of the different types of frames used in the multiple choice program is presented in Figure 3.

In the multiple choice program the subject was always faced with the task of selecting one of three response choices which was most closely related to the stimulus picture or word.

A number of variations were presented in the program, such as matching picture with picture, picture with word, word with word, and word with picture. In addition, spelling frames were included which required the subject to select the correct letter to fill in the missing letter of a word presented in the stimulus position. For example, the subject is required to complete "B__X" by pressing the key under the response selection "O".

The program also contained frames in which only the stimulus would appear on the first card, and the three response choices would appear on the following card. For example, a picture of a cat would be presented on the first card with the words "cat," "dog," "ball" appearing on the second card as response choices.

Throughout the programing, emphasis was placed upon immediate repetition of a new word. When a new word was introduced into the program, it appeared on successive cards for approximately eight frames. Thereafter, the word appeared intermittently with reduced frequency over five sessions. The word was then reintroduced every fifth session for the remainder of the program.

Keyboard Method

In his original method, Moore projected letters on the screen of the slide projector and requested that his subjects find these

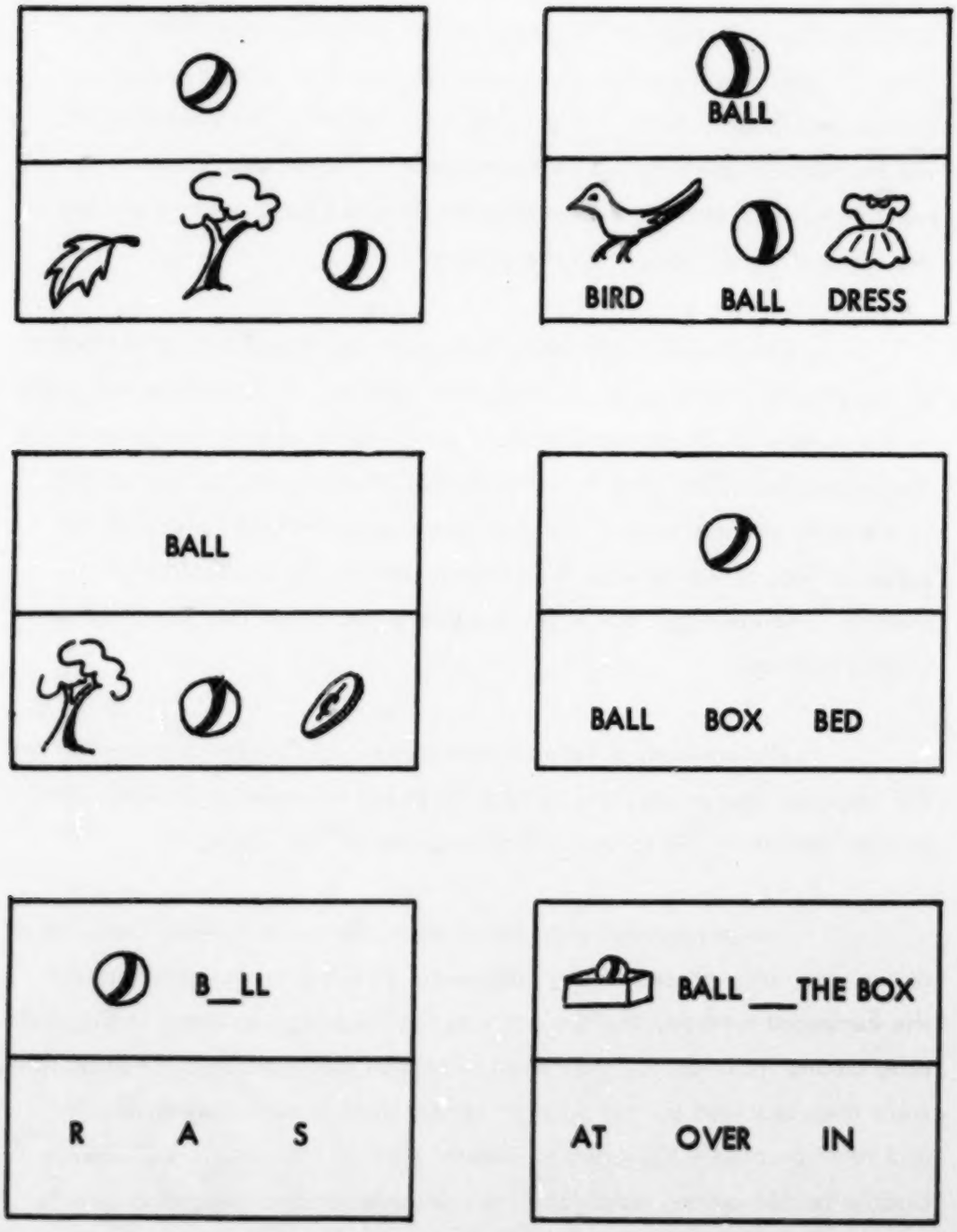

Fig. 3. Multiple Choice Program -- Sample Frames

letters on the typewriter keyboard and copy them. In the Wyckoff modification of the Moore technique, the typewriter mechanism was eliminated; and the action of the individual letters being typed is simulated by the prepared film strip which advances and shows the letter "typed" by the subject. Thus each frame for the keyboard method requires successful discrimination, location, and pressing of the individual key required by each frame. The subject's reinforcement consists of seeing and hearing the advance mechanism work and the letter "typed" appear on the screen.

Over 6,000 individual frames were prepared for presentation of the programed words by the keyboard method. This number was equal to the number of frame presentations received by clients on the multiple choice method. The programed words were introduced to the subject in the same sequence as in the multiple choice method, pictorial representations of words were the same as used in the multiple choice method, and the type face was the same as that used in the multiple choice method.

The methods differed in the actual mechanics of presentation, the response equipment, the number of possible response choices, and to some extent in the learning task required of the subject.

The perceptual tasks required of the subject were most similar during the early phases of the programs. In the program prepared for the keyboard method, the subject was first required to learn to discriminate among individually presented letters of the alphabet. Discriminations were then required among pairs of letters at first obviously dissimilar and later purposely designed to require finer discriminations. Simple discrimination among words was then introduced and later modified to include discriminations among words accompanied by their pictorial representations. In later stages of complexity, the word or its pictorial representation would be removed, thus requiring the subject to respond from memory.

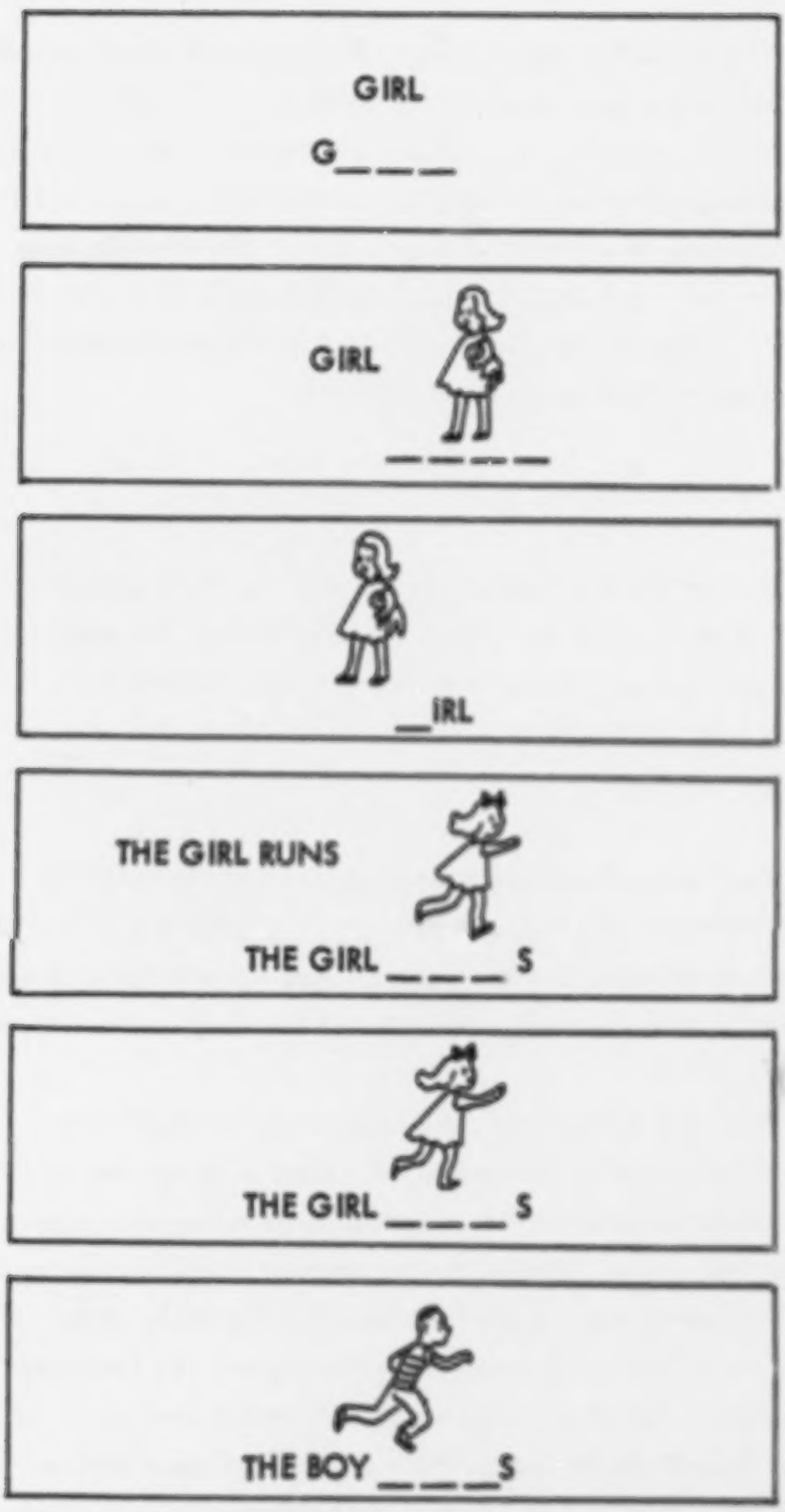

Fig. 4. Keyboard Method -- Sample Frames

The number of repetitions given a single word was maintained at approximately the same level as the multiple choice program. In the keyboard method, however, the subject was required upon every frame to supply missing letters of the word rather than to make simple discriminations among whole word configurations. While whole word configurations were required, single letter discriminations also were required for a response. Examples of the various types of frames used in the keyboard method are given in Figure 4.

Administration Conditions

During the data collection phase of the study special rooms were set up in each of the ten public schools in which the teaching machines were located. These rooms were usually classrooms or office space reserved for the project during the hours when the research was scheduled.

Four research assistants were assigned to the teaching machine instruction. All research assistants had professional training and previous experience in education but were not officially connected with the public school system at the time of this study.

Machine instruction lasted approximately eight weeks. Students worked at the machines daily during this period from fifteen to twenty minutes per day. The subjects followed a regular schedule for instruction, coming to the teaching machine room at an appointed time each day. Upon completion of the daily lesson, they returned to their regular classrooms. In most cases, two teaching machines were in operation in the same room at the same time. In some cases, this would be two multiple choice machines; and in other cases, this would be one keyboard machine and one multiple choice machine. Under non-experimental conditions it is estimated

that it would be possible for one adult to supervise the operation of between five and ten machines running concurrently.

For multiple choice machine instruction, subjects entered the room, picked up the cards to be used for that day from the research assistant, loaded the machines independently, and operated the machines with a minimum of supervision. The research assistant was available in case of technical difficulty arising from machine malfunction.

The same administration conditions prevailed for the keyboard method. Because of the complexity of the electronically operated keyboard apparatus, a higher degree of supervision was required by the research assistant. At times the film advance mechanism would require minor adjustments to properly align the frames in the viewing window. It was also necessary to assist some students in cases where prolonged response delay was encountered. This was encountered more frequently during the early lessons because the student was required to make one correct selection out of a possible 32 choices in contrast to the 1:3 response selection ratio required by the multiple choice machines.

Similar procedures were followed at the Sunland Training Center, with the exception that the instruction room was located in the hospital building; and students would come from their residence cottages for machine instruction on a regular daily schedule. One full time research assistant was employed to supervise all institutional subjects on both methods.

Tutorial Instruction

From the time the study was first formulated the investigators were aware of inherent potential for the "Hawthorne effect" in the use

of teaching machine instruction. In the latter part of the study it was decided to evaluate the novelty effect of machine instruction by selecting ten additional matched subjects for individual tutoring.

Two of the research assistants who had supervised machine instruction gave individual instruction in reading to the ten tutorial subjects. This instruction took place in the same schools at the same time that machine instruction was in process. Time allocated to each student was the same as for the machine instructional methods; thus overall the tutorial subjects received the same total instructional time as the machine instructed subjects.

Each research assistant giving instruction on a tutorial basis was a certified teacher. Kirk's "Teaching Reading in Slow Learning Children" (1940) was used as a guide to instructional content. Research assistants emphasized the same 72 programed words and 18 programed spelling words presented on the machine programs.

Stastical Treatment

In creating the subject triads the matching variables were considered in the following order of importance: sex, age, intelligence, grade placement as measured by the Gates Primary Reading Test, and words known among the 72 programed words. While it became evident later that knowledge of the 72 programed words predicted ability to profit from machine instruction, knowledge of the 72 programed words was only a subsidiary matching variable at the time subject triads were chosen. After matching, subjects were assigned randomly among the two experimental methods and the classroom method.

Pre and post-test measures are discussed earlier in this section. Gains scores are considered to be the difference between

pre-test scores and post-test scores on the several variables measured. Retention scores are expressed as ratios of the gains on periods from post-instruction test to the 30 day post-test and from the post-instruction test to the 60 day post-test.

Simple analysis of variance was run on all gains scores. Where significant F's were found, t test were run between treatment methods.

Point biseiral r's were run to determine relationships between sex and instructional methods. Product moment correlations were run to measure relationships between teaching procedures and the following variables: CA, MA, programed words known prior to special instruction, and pre-instruction to post-instruction word gains.

PRECEDING PAGE BLANK-NOT FILMED

CHAPTER III

RESULTS

Hypothesis 1 stated that no significant differences in efficiency will be found between automated teaching procedures and standard classroom instruction procedures for teaching spelling and reading to mental retardates. The major significant variables chosen to determine the efficacy of the two experimental procedures for teaching reading and spelling were the pre-test to post-test gains in ability to recognize the 72 programed words and the pre-test to post-test gain in ability to spell.[1]

In addition to the specific measures of gain in word recognition and spelling among the programed and non-programed words, it was thought desirable that pre and post-testing also be done utilizing equivalent forms of an already standardized instrument, the Word Recognition phase of the Gates Primary Reading Test (Gates, 1958). While scores on the Gates test are normally reported as grade equivalencies, Gates raw scores will be used unless otherwise indicated for the purposes of this study. The decision to use the raw scores was based upon the greater ease in statistical computation and because

---

[1] For the purposes of ease in discussion, the 72 words programed for this study will be referred to as the "programed words", and the 28 words of similar difficulty which make up the remainder of the 100 word list will be referred to as the "non-programed words." In like manner, the 18 words constituting 25 per cent of the programed words chosen as a measure representative of gains in spelling will be referred to as "programed spelling." The 7 words constituting 25 per cent of the non-programed words chosen as a measure of ability to spell among the non-programed words will be termed "non-programed spelling."

the positively skewed distribution due to the 1.3 basal grade level of the Gates test prevented adequate representation of the lower end of the sample.

The Public Schools Group

Table 3 provides a summary of performance among the experimental and control groups in terms of pre-test to post-test gain, pre-test to 30 day post-test gain, and pre-test to 60 day post-test gain on the measures cited above. The data shown in this table represent the most important findings of the study. Among the public schools subjects the mean pre-instruction to post-instruction word gain for the typewriter keyboard method was 28.10 words. For the multiple choice method, the gain was 20.50 words. The gains for these groups compare favorably to a mean word gain of only 3.70 for the normal classroom subjects. Analysis of individual subject data shows that, in every case, EMR pupils more than doubled the number of words known prior to automated instruction. While in some cases this gain consisted of only a few words, it should be noted that these instructional media made it possible, in approximately ten hours of instruction, to double the gains of from four to ten years of previous public school instruction.

The 30 day post-test revealed that for the multiple choice, keyboard and classroom groups respectively, means of 16.1 words, 19.7 words and 5.8 words were retained from those originally learned. The 60 day post-test showed that means of 18.2 words, 21.73 words and 5.8 words , respectively, were still retained.

During the pre-instruction to post-instruction period, gains on the 18 spelling words averaged 3.00 for the multiple choice group, 3.80 for the keyboard group, and 0.55 for the classroom group. Variance among these scores was high, and the gains were not well retained.

TABLE 3

WORD GAINS FOLLOWING INSTRUCTION

PUBLIC SCHOOLS SAMPLE

	Mean Gain *		
	Multiple Choice	Keyboard	Classroom
Programed Words (N=72)			
Pre-Test to Post-Test	20.50	28.10	3.70
Pre-Test to 30 Day Post-Test	16.10	19.70	5.80
Pre-Test to 60 Day Post-Test	18.20	21.73	5.80
Programed Spelling Words (N=18)			
Pre-Test to Post-Test	3.00	5.80	.55
Pre-Test to 30 Day Post-Test	1.20	2.70	.95
Non-Programed Words (N=28)			
Pre-Test to Post-Test	2.50	4.80	1.00
Pre-Test to 30 Day Post-Test	1.80	4.20	2.35
Pre-Test to 60 Day Post-Test	2.45	5.20	2.40
Non-Programed Spelling Words (N=7)			
Pre-Test to Post-Test	.55	1.10	.09
Pre-Test to 30 Day Post-Test	.65	.16	.23
Gates Word Recognition Test			
Pre-Test to Post-Test	.18	8.04	4.86
Pre-Test to 30 Day Post-Test	2.10	5.63	6.60

* Standard Deviations are given in Appendix F.

At the 30 day post-test, word gains for the multiple choice group were 1.2, for the keyboard group 2.7, and for the classroom group .9.

Pre-Test to post-test gains on the 2° non-programed words showed uniformly greater generalization for the keyboard group. Gains over this period were 2.5 words for the multiple choice group, 4.8 words for the keyboard group, and 1.0 word for the classroom group. The 30 day post-test showed gains of 1.8, 4.2, and 2.3 words for these groups respectively. The 60 day post-test showed some recovery for the multiple choice group with a 2.4 mean word gain, while a 5.2 mean word gain was shown by the keyboard group, and a 2.4 word gain was shown by the classroom group.

Although the 7 non-programed spelling words constitute a very small sample, results are reported for the sake of uniformity. Pre-test to post-test gains showed means of .5 words for the multiple choice group, 1.1 words for the keyboard group, and .09 words for the classroom group. Pre-test to 30 day post-test gains were .65, .16, and .23 words for each group respectively.

On the Word Recognition section of the Gates Primary Reading Test, the mean raw score gain for the multiple choice method was negligible (.18 words). On the other hand, the keyboard method showed a sizeable raw score gain with 8.04 words for the pre-instructional to post-instructional period. Interestingly, the classroom group also showed a high relative gain with a mean of 4.9 words. At the 30 day post-testing, scores for the multiple choice group had risen with a mean of 2.1 words as opposed to means of 5.6 words for the keyboard group, and 6.6 words for the classroom group.

The Institutional Group

Thirty-three institutional subjects from Sunland Training Center, Ft. Myers, Florida, were grouped as 11 triads of matched subjects. The only available evaluation measures for these subjects were the programed and non-programed words lists and the Gates Primary Reading Test. In other respects the procedure followed with these subjects paralleled the procedure followed with the public schools subjects.

Certain differences exist between the institutional subjects and the public schools subjects. The descriptive data presented in Table 2 shows that the institutional subjects were three years older, on the average, than the public schools subjects and were slightly higher in mental age. Table 4 provides a summary of performance among the experimental and classroom groups for this sample. While word gains by the institutional group were impressive, they were not of the same order as those for the public schools group. The ratio of increase in words known, however, remained relatively constant with the greater portion of the institutional group doubling pre-test word recognition scores.

As with the public schools group, individual scores frequently showed what appeared to be a reminiscence effect on delayed post-testing.[2] With the institutional sample, mean gains on the programed words over the pre-instructional to post-instructional period were 15.7 words for the multiple choice group, 14.5 words for the keyboard group, and .8 words for the classroom group. At the end of 30 days,

---

[2]See Appendices D and E, for example, multiple choice group subjects 7 and 19 or keyboard group subjects 1 and 15.

TABLE 4

WORD GAINS FOLLOWING INSTRUCTION

SUNLAND TRAINING CENTER

	Mean Gain *		
	Multiple Choice	Keyboard	Classroom
Programed Words (N=72)			
Pre-Test to Post-Test	15.70	14.50	.80
Pre-Test to 30 Day Post-Test	15.30	11.40	2.80
Programed Spelling Words (N=18)			
Pre-Test to Post-Test	7.80	9.63	-
Pre-Test to 30 Day Post-Test	7.00	7.09	-
Non-Programed Words (N=28)			
Pre-Test to Post-Test	1.27	1.18	-.18
Pre-Test to 30 Day Post-Test	1.63	1.63	.55
Non-Programed Spelling Words (N=7)			
Pre-Test to Post-Test	1.09	.82	-
Pre-Test to 30 Day Post-Test	1.30	1.09	-
Gates Word Recognition Test			
Pre-Test to Post-Test	2.36	2.90	2.36
Pre-Test to 30 Day Post-Test	1.80	4.09	1.70

*Standard Deviations are given in Appendix G.

gains for the multiple choice group remained approximately the same with 15.3 words. Scores had decreased for the keyboard group with 11.4 words, and had increased for the classroom group with 2.8 words.

Proportional gains on the programed spelling words were higher for the institutional group than for the public schools group, with mean pre-test to post-test gains of 7.8 words for the multiple choice method and 9.6 words for the keyboard method. At the end of 30 days, spelling word gains remained high, with a 7.1 word gain for the multiple choice method and a 7.0 word gain for the keyboard method. Post-testing on spelling with the classroom group was not done.

When compared to the public schools group, the institutionalized subjects had approximately the same relative gains among the non-programed words, although variance was considerably greater. On post-testing the mean gain for the non-programed words for the multiple choice group was 1.3, for the keyboard group was 1.2, and for the classroom group was -.2. At the end of 30 days, scores for all three groups were slightly higher, with 1.6 word gains for both the multiple choice and keyboard groups, and a .6 word gain for the classroom group.

Proportional gains on the non-programed spelling words were approximately the same for the institutional subjects as for the subjects from the public schools, with pre-test to post-test gains of 1.1 words for the multiple choice group and .8 words for the keyboard group. At the 30 day post-testing, gains were 1.3 words for the multiple choice group and 1.1 words for the keyboard group respectively. No measure of non-programed spelling was done among the institutional classroom group.

The Tutorial Group

Among the advantages cited for programed instruction are individualized instruction and the immediate feedback of information. The question has been posed whether such feedback would be superior if provided by the teacher working on the one-to-one basis with a student. Because of this question and because of the inherent potential for the Hawthorne effect in the singling out of students for machine instruction, a special group was selected so that the keyboard and the multiple choice methods could be compared with direct tutorial instruction of the individual student. Ten new subjects were drawn for the tutorial group and were matched with ten subjects from the multiple choice group, from the keyboard group, and from the classroom group. Gains of the tutorial subjects are compared to the multiple choice, keyboard, and classroom groups in Table 5.

Over the pre-instructional to post-instructional period, gains were uniformly higher for the tutorial group on the programed words, on the non-programed words and on the Gates Primary Reading Test, Word Recognition section. The keyboard group made the highest gains on the programed spelling words and on the non-programed spelling words.

For the programed words over the pre-instructional to post-instructional period, the tutorial group gain was 32.3 words compared to 21.4 words for the multiple choice group, 30.1 words for the keyboard group and 2.3 words for the classroom group.

For the programed spelling words, pre-test to post-test scores showed the tutorial group gaining only 1.7 words as opposed to 3.2 for the multiple choice group, 7.3 for the keyboard group, and .2 for the classroom group.

TABLE 5

COMPARISON OF MATCHED GROUP PERFORMANCE

MEAN WORD GAINS

	Multiple Choice (N=10)	Keyboard (N=10)	Classroom (N=10)	Tutorial (N=10)
	Mean Gain	Mean Gain	Mean Gain	Mean Gain
Programed Words				
Pre-Test to Post-Test	21.4	30.1	2.3	32.3
Pre-Test to 30 Day Post-Test	15.5	22.3	4.2	18.7
Programed Spelling Words				
Pre-Test to Post-Test	3.2	7.3	.2	1.7
Non-Programed Words				
Pre-Test to Post-Test	2.2	7.1	1.1	8.5
Pre-Test to 30 Day Post-Test	1.4	6.2	2.0	3.5
Non-Programed Spelling Words				
Pre-Test to Post-Test	.7	1.5	.1	.5
Gates Word Recognition Test				
Pre-Test to Post-Test	.2	7.7	4.2	11.0

Standard Deviations are given in Appendix J.

The tutorial group appeared to have greater generalization than any of the experimental groups. On the non-programed words, this group showed a mean gain over the pre-test to post-test period of 8.3 words as opposed to 2.2 words for the multiple choice group, 7.1 words for the keyboard group, and 1.1 words for the classroom group.

The keyboard group continued to show the highest word gain on the non-programed spelling words with a 1.5 word gain over the pre-test to post-test period. The tutorial group showed a .5 word gain, the multiple choice group a .7 word gain, and there was a .08 gain for the classroom group.

On the Gates Primary Reading Test, Word Recognition section, the tutorial group outstripped the others with a mean raw score gain of 11 points as opposed to gains of .2 for the multiple choice group, 7.7 for the keyboard group, and 4.2 for the classroom group.

Thirty day post-testing was done only on programed and non-programed words. The results, however, show relatively greater losses among the tutorial subjects than among the machine group subjects. Results over the pre-test to 30 day post-test period showed an 18.7 mean word gain score for the tutorial group as opposed to 15.5 words for the multiple choice group, 22.3 for the keyboard group, and 4.2 for the classroom group. A similar situation was obtained for the non-programed words, with mean word gains at the end of pre-test to post-testing of 3.5 for the tutorial group, 1.4 for the multiple choice group, 6.2 for the keyboard group, and 2.0 for the classroom group.

The ratio of pre-test to 30 day post-test retention is shown in Table 6. Retention of the tutorial group was considerably less than either of the experimental groups. On the programed words, the tutorial group had a .58 retention ratio as compared to .72 for the

TABLE 6

RETENTION RATIOS BETWEEN PRE-TEST TO POST-TEST GAINS AND PRE-TEST TO 30 DAY POST-TEST GAINS FOR THE SPECIAL MATCHED GROUPS

	Multiple Choice	Key-board	Class-room	Tutorial
Programed Words	.72	.74	1.8	.58
Non-Programed Words	.64	.87	1.8	.41

multiple choice group, .74 for the keyboard group, and 1.8 for the classroom group. On the non-programed words, retention ratio was .41 for the tutorial group as opposed to .64 for the multiple choice group, .87 for the keyboard group, and 1.8 for the classroom group. The relatively larger retention ratios of the classroom group are a function of very low word gains over the pre-test to post-test period.

Analysis of Retention

Hypothesis 2 states that effective retention of word recognition, reading and spelling skills obtained from automated instruction procedures will be demonstrated. Tables 3 and 4 show some losses in word recognition and spelling on 30-day and 60-day post-instruction tests in both samples for both experimental procedures. However, high percentage of retention was observed for word recognition, the Gates Test words, and non-programed spelling words; moderate retention of programed word spelling was also observed.

Hypothesis 3 states in part that no significant differences in retention will be observed between the two automated teaching procedures. Public school subjects taught by the keyboard method showed a higher mean number of words retained (post-instruction to 30-day and 60-day post-test scores) than the multiple choice group. However, the latter group showed higher percentages of words retained over these periods.

Analysis of variance was run on all measures representing retention among the three instruction groups. F ratios are summarized in Appendix H.

For both the public school and institutional samples, no significant differences were observed in retention on any of the measures.

The Programed Words

For the public schools group, simple analysis of variance showed that gains in word recognition were significant at the .01 level for the pre-test to post-test, for the pre-test to 30 day post-test, and for the pre-test to the 60 day post-test periods.

A t test was employed to evaluate differences between the three possible pairs of treatment groups over the pre-test to post-test periods. A summary table of all t ratios is given in Appendix I.

Mean performance scores of the keyboard method and of the multiple choice method differed significantly from mean performance scores of the classroom method at the .01 level of confidence for the pre-test to immediate post-test period. Mean performance scores on

the multiple choice method also differed from mean performance scores on the keyboard method at the .05 level of confidence with higher gains favoring the keyboard method.

Testing of mean gains over the pre-test to 30 day post-test period revealed that the multiple choice method and the keyboard method did not differ significantly from each other, although both methods continued to differ from the classroom method at the .01 level of confidence.

Similarly, gains scores from the pre-test to 60 day post-test differed significantly from mean classroom gains, while the keyboard and multiple choice methods were not significantly different from each other.

With the institutional group, gains in the number of programed words known were measured at post-test and 30 day post-test intervals only. Gains scores over the pre-instructional to post-instructional period and over the pre-instructional to the 30 day post-instructional period yielded $\underline{F}$ ratios significant at the .01 level.

The $\underline{t}$ tests showed results paralleling results found in the public schools sample, with the multiple choice and the keyboard groups differing at the .01 level of confidence from the classroom groups. The multiple choice group differed from the keyboard group at the .05 level of confidence with results favoring the keyboard method, although this difference disappeared in comparisons of gain over the pre-instructional to 30 day post-instructional period. At 30 day post-testing, the keyboard method group differed from the classroom group at the .01 level of confidence; while the multiple choice method group differed from the classroom group at the .05 level of confidence.

Among the special group of subjects matched for the inclusion of tutorial subjects, the *F* ratio was significant at the .01 level; and a *t* ratio showed all three treatment groups differing significantly at the .01 level of confidence from the classroom group. The tutorial, multiple choice, and keyboard groups did not differ significantly from each other; although mean gains for the tutorial group were higher than either the keyboard or multiple choice method at post-testing. Over the period from pre-testing to 30 day post-testing the tutorial group, along with the keyboard and multiple choice groups, continued to show significant differences from the classroom group but did not differ significantly from each other.

The Programed Spelling Words

On the programed spelling words, the public schools subjects yielded an *F* ratio significant at the .01 level among the mean group gains over the pre-test to post-test period. Pre-test to 30 day post-test gains on the programed spelling words did not yield a significant *F*. Comparison of individual group means by *t* ratio showed differences significant at the .01 level between the keyboard and the classroom groups and between the multiple choice and the classroom groups. The multiple choice and the keyboard groups were not significantly different from each other.

For the institutional group, *t* ratios between the multiple choice and the keyboard groups programed spelling gains were not significant either at post-testing or at 30 day post-testing.

For the special group of subjects matched for the inclusion of tutorial subjects, an *F* ratio significant at the .01 level was obtained for the pre-test to post-test period. The *t* ratios between the five possible inter-group combinations over this period showed significance

at the .05 level between the tutorial and the keyboard methods and between the multiple choice and the classroom methods. Significance at the .01 level was found between the classroom group and the keyboard group. In each case, the difference favored the keyboard group.

The Non-Programed Words

Single analysis of variance of the non-programed word gain for the public schools students yielded an F ratio significant at the .01 level on gains over the pre-test to post-test period. From the pre-test to the 30 day post-test and from the pre-test to the 60 day post-test, F ratios were not significant.

Analysis of means by t ratios showed that for the public schools group the keyboard method differed significantly, at the .01 level, from the classroom group. The keyboard group did not differ significantly from the multiple choice group, and the multiple choice group did not differ significantly from the classroom group.

For the institutional subjects, no significant F ratios were found for the non-programed words.

For the special group of subjects matched for the inclusion of tutorial subjects, an F ratio significant at the .05 level was found on gains from pre-testing to post-testing. Over the period from pre-testing to 30 day post-testing, the F ratio was not significant.

For the pre-test to post-test period, analysis of t ratios among the five possible inter-group combinations showed a .05 level significance only between the keyboard group and the classroom group and between the tutorial group and the classroom group. All other ratios were not significant.

The Non-Programed Spelling Words

For the public schools subjects, an *F* ratio calculated among group means on the non-programed spelling words was significant at the .01 level for the pre-test to post-test period gains, but was not significant among gains over the pre-test to 30 day post-test period. The *t* ratios showed significant differences only between the keyboard group and the classroom group. No significant difference was found between the keyboard and the multiple choice groups or between the multiple choice and the classroom groups.

Among the institutional subjects, analysis of *t* ratios between the keyboard method and the multiple choice method on gains over the pre-test to post-test and over the pre-test to 30 day post-test periods yielded no significant differences.

For the special group of subjects matched for the inclusion of tutorial subjects, the *F* ratio for the non-programed spelling words was not significant.

The Gates Primary Reading Test - Word Recognition

Analysis of raw scores on the word recognition section of the Gates Primary Reading Test for public schools subjects revealed an *F* ratio significant at the .01 level among gains over the pre-test to post-test period. The *F* ratio for mean gains over the pre-test to 30 day post-test period for this group was not significant.

Interestingly, the multiple choice and the keyboard methods were significantly different from each other at the .01 level of confidence on this measure with results favoring the keyboard method. No significant differences were found between the keyboard group and the

classroom group. However, mean gains of the multiple choice group differed from the classroom group at a .05 level of confidence with the classroom group showing the greater score increase.

F ratios calculated among the institutional sample groups over the pre-test to post-test instruction and over the pre-test to 30 day post-test instruction were not significant.

The special group of subjects matched for the inclusion of tutorial subjects, yielded a significant *F*. Analysis of the group inter-relations showed only one significant *t* ratio, between the multiple choice group and the tutorial group with results favoring the tutorial group. No significant differences were found between the other groups.

Post-Instruction Retention

Analysis of retention heretofore has been in terms of pre-instruction to 30-day and 60-day post-instruction scores. Retention was also analyzed in terms of differences between post-instruction to 30-day and 60-day post-instruction scores.

Compared to the multiple choice group, in the public school sample, the keyboard group began with a higher number of words known (Table 1), learned more words during automated instruction and retained a greater number for 30 and 60 days afterwards (Table 3). However, the multiple choice group retained a higher percentage of words than did the keyboard group. *Ss* taught by the multiple choice procedure retained 78.6% of words learned 30 days after instruction; this increased to 88.8% retention on the 60-day post-test. The keyboard group also retained very high proportions of words learned under this method; namely, 66.6% and 77.4%, respectively, on the 30-day and 60-day post-tests. The

same trend was observed for retention of programed word-recognition in the institutional sample.

However, retention of programed spelling words over a 30-day post-instruction period was more modest for both experimental groups. In the public school sample the multiple choice group retained 40% of the spelling words learned while the keyboard group retained 47% of spelling words learned during automated instruction. Institutionalized retardates both learned to spell and retained a higher percentage of words under both experimental procedures (Table 4). The multiple choice group retained 89% and the keyboard group 72% of spelling words learned during automated instruction.

Retention effects were negligible in terms of numbers of words involved for non-programed word recognition, non-programed spelling and the Gates Word Recognition Test. Consequently, analyses of retention percentages for these measures are not included here. They can be estimated from data provided in Tables 3 and 4.

Relationships Between Instructional Condition & Other Variables

Product moment correlations were calculated to determine if any of the variables on which subjects were matched were related to ability to profit from teaching machine instruction. Since the major variable showing ability to profit from teaching machine instruction was considered to be the number of programed words learned over the pre-instructional to post-instructional period, these word gains were correlated with sex, chronological age, and mental age. It was observed in the pilot study that the number of words known by the subject on pre-testing appeared to be related to the individual subject's ability to profit from instruction. Consequently,

the number of programed words known prior to instruction was included in the correlational analysis.

Correlation With Programed Word Gain

Relationships between the programed word gain and other variables for the public schools group are shown in Table 7.

TABLE 7

RELATIONSHIPS BETWEEN PROGRAMED WORD GAIN & RELATED VARIABLES

A. PUBLIC SCHOOL GROUPS

	PROGRAMED WORD GAIN		
	Multiple Choice	Keyboard	Classroom
Sex	.05	.43	.18
Chronological Age	.44	.37	.54
Mental Age	.20	.03	.29
Programed Words Known on Pre-Test	.80	.72	.65

Sex showed no appreciable correlation to programed word gain for the multiple choice group and the classroom group. It did, however, show moderate correlation with performance of the keyboard group. The male subjects appeared generally to gain from instruction more readily than female subjects. All three groups showed moderate correlation with chronological age. For chronological age, Pearson r's were .44 for the multiple choice group, .37 for the keyboard group, and .54 for the